

CLAIM PTO

10/02/00

T.D.

1. A method of transmitting data from a source file located at a sending computer to a receiving computer, the computers being connected through a computer data interface, the method comprising the steps of:

(a) dividing a reference file located at the receiving computer into a plurality of data blocks, each data

block having a length of n bytes, and associating each data block with a reference key value determined in accordance with a key defining method by the data in that block:

(b) transmitting the reference key values to the sending computer:

(c) identifying blocks of data of length n bytes from the source file, determining source key values in accordance with the key defining method, and using the source and reference key values to compare blocks of data from the reference file to blocks of data from the source file and, in instances where a match is found between a block of data from each file, sending an indication of the match to the receiving computer so that the block of data indicated by the match need not be transmitted to the receiving computer.

2. The method of claim 1 wherein an initial block of data is identified from the source file, a source key is determined from the initial block, and if a match for the initial block is not found:

the method further including the step of transmitting a subset of the initial block to the receiving computer, the subset including less than all of the information in the initial block; and

identifying blocks of data of length n bytes from the source file includes identifying from the source file a subsequent block of data of length n bytes comprising the initial block of data, less the transmitted subset, and additional data from the source file.

3. The method of claim 1 wherein at least a portion of the key value for a block of data is computed by adding the value of each byte of data in the block to produce a total of all of the bytes in the block.

4. The method of claim 1 wherein at least a portion of the key value for a block of data is computed by multiplying the bytes in the block by one or more multipliers, the value of the multiplier being dependent upon the position of a given byte in the block, and summing the results of the multiplication operations.

5. A method of transmitting data from a source file located at a sending computer to a destination file located at a receiving computer, the computers being connected through a computer data interface, the method comprising the steps of:

- (a) identifying a reference file at the receiving computer that may have data identical to some of the data comprising the source file;
- (b) dividing the data comprising the reference file into a plurality of data blocks having n-bytes per block and associating each data block with a reference key value determined by a key defining method;
- (c) identifying an n-byte block of data from the source file and computing using the key defining method a current value for a source key associated with the identified block of data;

- (d) comparing the current value of the source key with each of the reference key values and, if a match is found, (i) transferring an indication of such to the receiving computer, and (ii) repeating step (c); and
- (e) if a match was not found in step (d), transferring to the receiving computer a subset including less than all the data in the n-byte block of data, removing the subset from the n-byte block of data, adding additional data from the source file to the n-byte block of data, re-computing using the key defining method a current value of the source key, and repeating step (d).

6. The method of claim 5, wherein steps (c) and (d) are repeated only until all of the data in the source file has been considered.

7. (Amended) The method of claim 5 wherein recomputing a current value of the source key of step (e) includes deriving at least a part of the current value of the source key from at least a part of the previous source key by removing the contribution to the part of the source key from the transmitted subset and integrating into the part of the source key a contribution from the additional data.

8. The method of claim 1 wherein each byte of data in the reference file is used in the determination of not more than one reference key, and in which at least some of the bytes of data in the source file are used in the determination of multiple source keys.

9. The method of claim 2 in which the source key and the reference key include multiple bits and in which some of the bits are determined by a summing operation and some of the bits are determined by a logical operation.

10. The method of claim 9 in which the summing operation includes multiplying by constant coefficients the values represented by bytes of the blocks of source data and in which the logical operation comprises an exclusive OR operation.

11. The method of claim 2 wherein the key defining method for the blocks of data includes the following calculation:

$$C_1(n) + C_2(n-1) + C_3(n-2) + \dots + C_{n-1}(2) + C_n$$

where C_n is the character in the n th position of the block of data.

12. The method of claim 11 wherein the key defining method includes the following logical operation:

$$C_1 \text{ XOR } C_2 \text{ XOR } C_3 \dots C_{n-1} \text{ XOR } C_n$$

13. A method of changing data at a receiving unit until the data at the receiving unit is identical to data at a source unit, comprising:

- (a) determining multiple reference keys corresponding to groups of data stored at the receiving unit;
- (b) transmitting the multiple reference keys to the source unit;
- (c) determining a source key corresponding to a group of source data in the source unit;
- (d) comparing the source key with the multiple reference keys;
- (e) transmitting data from the source unit to the receiving unit if the source key does not match any of the reference keys;
- (f) transmitting a control signal from the source unit to the receiving unit if the source key matches a reference key, the control signal causing the receiving unit to use data at the receiving unit corresponding to the matched reference key; and
- (g) repeating steps (c), (d), (e), and (f) for additional groups of source data in the source unit until the data at the receiving unit is identical to the data at the source unit.

14. (Amended) The method of claim 13 wherein the data transmitted is a subset of the group of source data associated with the matching source key[, the subset including less than all of the information in the initial block].

15. (Amended) The method of claim 13 wherein each byte of data in the [reference file] receiving unit is used in the determination of not more than one reference key, and in which at least some [of the] bytes of the data in the source [file] unit are used in the determination of multiple source keys.

16. An apparatus for changing data at a receiving unit until the data at the receiving unit is identical to data at a source unit, comprising:

means for determining an array of reference keys corresponding to groups of data stored at the receiving unit;
data transfer means for transmitting the multiple reference keys to the source unit;

means for determining source keys corresponding to groups of source data in the source unit;

means for comparing the source keys with the multiple reference keys;

means for transmitting data from the source unit to the receiving unit when a source key does not match any of the reference keys; and

means for transmitting a control signal from the source unit to the receiving unit when a source key matches a reference key, the control signal causing the receiving unit to use a group of data at the receiving unit corresponding to the matched reference key.

17. (Amended) The apparatus of claim 16 wherein the means for determining source keys determines a new source key after the means for comparing source keys has compared the previously determined source key, and wherein the means for determining source keys determines the new source key from a group of source data, the composition of [the] which is determined by whether the previously compared source key matched a reference key.

18. (Amended) A method of creating at a receiving computer a duplicate file that is identical to a source file at a sending computer, the duplicate file being formed in part from data in a reference file located at the [remote] receiving computer and in part from data in the source file transmitted from the sending computer, the computers being connected through a computer data interface, the method comprising the steps of:

(a) dividing the reference file located at the receiving computer into a plurality of data blocks of uniform length and associating with each data block a reference key value determined by the data in that block in accordance with a key defining method; and

(b) identifying blocks of data of the uniform length from the source file, determining source key values in accordance with the key defining method, and comparing the source and reference key values to determine whether blocks of data from the reference file match blocks of data from the source file and, in instances where a match is found between a block of data from each file, sending an indication to the receiving computer to copy the block of data from the reference file into the duplicate file so that the block of data indicated by the match need not be transmitted to the receiving computer, wherein the blocks of data from the source file are sequentially identified and each source block of data includes some of the data from the preceding source block of data if the preceding source block of data did not match a reference block of data.

19. The method of claim 18 in which the uniform length of the data blocks is at least 256 bytes.

20. The method of claim 18 in which the key defining method defines keys that are at least 32 bits in length.

21. (Amended) A method of creating at a receiving computer a duplicate file that is identical to a source file at a sending computer, the duplicate file being formed in part from data in a reference file located at the [remote] receiving computer and in part from data in the source file transmitted from the sending computer, the computers being connected through a computer data interface, the method comprising the steps of:

(a) dividing the reference file located at the receiving computer into a plurality of data blocks of uniform length and associating each data block with a reference key value determined by the data in that block in accordance with a key defining method;

(b) identifying blocks of data of the uniform length from the source file, determining source key values in accordance with the key defining method, and using the source and reference key values to compare blocks of data from the reference file with blocks of data from the source file;

(c) in instances where a match is found between a block of data from each file, sending an indication of the match to the receiving computer to copy the block of data from the reference file to the duplicate file so that the block of data indicated by the match need not be transmitted to the receiving computer; and

(d) in instances where a match is not found between a block of data from each file, transmitting [less] fewer bytes than the number of bytes in the uniform length from the source file to the receiving computer and adding transmitted bytes to the duplicate file.

22. (Amended) A method of creating at a receiving computer a duplicate file that is identical to a source file at a sending computer, the duplicate file being formed in part from data in a reference file located at the [remote] receiving computer and in part from data in the source file transmitted from the sending computer, the computers being connected through a computer data interface, the method comprising the steps of:

(a) dividing the reference file located at the receiving computer into a plurality of data blocks, each data block having a length of n bytes, and associating each data block with a reference key value determined by the data in that block in accordance with a key defining method;

(b) identifying blocks of data of length n bytes from the source file, determining source key values in accordance with the key defining method, and using the source and reference key values to compare blocks of data from the source file with blocks of data from the reference file to find a match;

(c) in instances where a match is found between a block of data from each file, sending an indication of the match to the receiving computer to copy the block of data from the reference file to the duplicate file so that the block of data indicated by the match need not be transmitted from the [source] sending computer to the receiving computer; and

(d) in instances where a match is not found:

(i) transmitting a subset of [the] an initial block to the receiving computer and adding the subset to the duplicate file;

(ii) identifying from the source file a subsequent block of data of length n bytes comprising the initial block of data, less the transmitted subset, and additional data from the source file; and

(iii) determining for the subsequent block of data a source key, the source key being derived from the source key determined from the initial block of data by removing the contribution from the transmitted subset and incorporating a contribution from the additional data.

23. The method of claim 13 wherein if the preceding source key did not match a reference key, the subsequent

source key is determined from a group of source data that includes some but not all of the data in the preceding group of source data and also includes data not included in the preceding group of source data and, if the preceding source key did match a reference key, the source key corresponds to a block of source data that directly follows the data used to determine the previous source key.

24. A method of changing data at a receiving unit until the data at the receiving unit is identical to data at a source unit, comprising:

- (a) determining using a key defining method multiple reference keys corresponding to data groups of length n bytes stored at the receiving unit;
- (b) transmitting the multiple reference keys to the source unit;
- (c) determining using the key determining method a source key corresponding to a group of source data of length n bytes in the source unit;
- (d) comparing the source key with the multiple reference keys;
- (e) transmitting data from the source unit to the receiving unit if the source key does not match any of the reference keys;

(f) transmitting a control signal from the source unit to the receiving unit if the source key matches a reference key, the control signal causing the receiving unit to use data at the receiving unit corresponding to the matched reference key; and

(g) repeating steps (c), (d), (e), and (f) for additional groups of source data in the source unit until the data at the receiving unit is identical to the data at the source unit.

wherein the groups of source data comprise, if the preceding source key did not match a reference key, $n-1$ bytes from the first group of data and one additional byte of data, and if the preceding source key did match a reference key, n bytes of data different from the n bytes of the preceding source group of data.

25. (Amended) An apparatus for changing data at a receiving unit so that the data at the receiving unit is identical to data at a source unit, comprising:
means for determining using a key defining method an array of reference keys having lengths [of] corresponding to data groups having a uniform length of at 256 bytes and stored at the receiving unit; data transfer means for transmitting the multiple reference keys to the source unit; means for determining using the key defining method source keys corresponding to groups of source data of the uniform length in the source unit; means for comparing the source keys with the multiple reference keys; means for transmitting, when a source key does not match any of the reference keys, less than all the data that is included in the group of source data used to determine the source key; and means for transmitting a control signal from the source unit to the receiving unit when a source key matches a reference key, the control signal causing the receiving unit to use a group of data at the receiving unit corresponding to the matched reference key.

26. A method of creating at a receiving computer a duplicate file that is identical to a source file at a sending computer, the duplicate file being formed in part from data in a reference file located at the receiving computer and in

part from data in the source file transmitted from the sending computer, the computers being connected through a computer data interface, the method comprising the steps of:

- (a) identifying a reference file at the receiving computer that may have data identical to the data comprising the source file;
- (b) dividing the data comprising the reference file into a plurality of data blocks having n-bytes per block and associating each data block with a reference key value determined by a key defining method;
- (c) transmitting the reference key values from the receiving computer to the sending computer;
- (d) identifying an n-byte block of data from the source file and computing using the key defining method a current value for a source key associated with the identified block of data;
- (e) comparing the current value of the source key with each of the reference key values and, if a match is found, (i) transmitting an indication of such to the receiving computer, which adds the matching data from the reference file to the duplicate file, and (ii) repeating step (d); and

- (f) if a match was not found in step (e), transferring to the receiving computer a subset of the n-byte block of data to be added to the duplicate file and repeating step (d).

27. The method of claim 26 in which, if a match was not found in step (d), a new block of data in the source file is defined by removing the transmitted subset from the previous n-byte block of data, adding additional data from the source file to the new n-byte block of data, re-computing using the key defining method a current value of the source key, and repeating step (e).

28. (Amended) The method of claim 26 in which, if a match is found in step (d), a new block of data in the source file is defined by the [n bytes] n-bytes immediately following the n-bytes used to form the previous source block of data.

29. (Amended) The method of claim 26 in which the reference keys comprise [and] a first part and a second part, the calculation of each part being independent of the calculation of the other part.

30. The method of claim 26 in which the uniform length of the data blocks is at least 256 bytes.

31. The method of claim 26 in which the key defining method defines keys that are at least 32 bits in length.

32. (Amended) A method of [making] creating a [first] reference data file at a first location that is identical to a source data file at a second location, the method comprising the steps of:

(a) identifying a reference file at the first location that may have data identical to the data comprising a source data file;

(b) dividing the data comprising the reference file into a plurality of data blocks having n-bytes per block and associating each data block with a reference key value determined by a key defining method;

(c) transmitting the reference key values from the first location to the second location;

(d) identifying an n-byte block of data from the source data file and computing using the key defining method a current value for a source key associated with the identified block of data;

(e) comparing the current value of the source key with each of the reference key values and, if a match is found, (i) transferring an indication of such to the [receiving computer] first location, which adds the matching data from the reference file to [the] a duplicate file, and (ii) repeating step (d); and

(f) if a match was not found in step (e), transferring to the [receiving computer] first location a subset of the n-byte block of data to be added to the duplicate file and repeating step (d).

33. A method of changing data stored at a receiving unit to match data stored at a source unit, comprising:

determining multiple first keys corresponding to groups of data stored at a first unit;

determining a second key corresponding to a group of data stored at a second unit;

comparing the second key with the multiple first keys;

designating which of the first unit and the second unit are the source unit and the receiving unit; and

transmitting from the designated source unit to the designated receiving unit data corresponding to the second key if the second key matches none of the multiple first keys, and leaving unchanged in the designated receiving unit the data corresponding to the second key if the second key matches one of the multiple first keys.

34. The method of claim 33 wherein the comparing the second key with the multiple first keys takes place in the second unit, and further comprising transmitting the multiple first keys from the first unit to the second unit.

35. The method of claim 34 wherein the first unit is the designated receiving unit.

36. The method of claim 33, further comprising:

successively determining different second keys corresponding to different groups of data stored at the second unit;

successively comparing each of the different second keys with the multiple first keys; and

transmitting from the designated source unit to the designated receiving unit data corresponding to each of the different second keys that matched none of the multiple first keys.

37. A method of transmitting data from a source computer to a receiving computer, the source and receiving computers being connected through a computer data interface, comprising:

dividing a first file into multiple data blocks and associating each data block of the multiple data blocks with a first key value determined in accordance with a key defining method by the data in the data block;

identifying multiple data blocks from a second file and determining second key values in accordance with the key defining method;

using the first and second key values to compare data blocks from the first file and from the second file;

designating which of the first file and the second file are located at the source computer and at the receiving computer; and

for instances in which a match is found between a data block from the first file and a data block from the second file, leaving unchanged the data block stored in the designated receiving computer.

38. The method of claim 37 wherein a selected data block from the first file is identified and a selected first key from the selected data block is determined and, for instances in which no match is found between a data block from the first file and a data block from the second file, the method further comprising:

transmitting to the designated receiving computer a subset of the selected data block from the first file, the subset including less than all of the information in the selected data block; and

identifying from the first file a subsequent data block comprising the selected data block less the subset transmitted to the designated receiving computer, and additional data from the first file.

39. The method of claim 37 wherein each of the multiple data blocks from the first and second files includes multiple bytes of data of which each byte has a value, and wherein at least a portion of the key value for a data block from any one of the first and second files is computed by adding the value of each byte of data in the data block to produce a total for all of the bytes in the data block.